

Partial English Translation of
LAID OPEN unexamined
JAPANESE PATENT APPLICATION
Publication No. 2000-253056

[0075] The request band input part 14 receives request band information specifying data quantity, which may be transmitted in time, from outside of the transmission band controller 30, and notifies it to the transmission band instructing part 16. The request band information is described in a form such as data quantity which the data transmission part 14 can transmit per second, or a transmission cycle in a minimum data unit of the data transmitted on the network 1. The request band information herein may have been statically stored since start-up of the system.

[0076] According to a band informed from the request band input part 15, the transmission band instructing part 16 calculates transfer data length per time, a cycle of data transfer and the number of repetition based on the following [Formula 1], and notifies it to the data transfer part 13 of the calculated result.

[0077]

[Formula 1] $B \text{ [Mbps]} = L \text{ [bits]} \times 1000000 / R \text{ [\mu s]}$

[0078] Where, B is a request band, L is data quantity which the data transfer part 13 transfers at a time, and R is a cycle which the data transfer part 13 transfers data. The request band is expressed in a Mbps unit.

[0079] For example, when the quantity of data transferred at a time is 64 bits and the band of 64 Mbps is requested, the transfer cycle is 1 micro second.

[0080] In addition, when a packet is transmitted to the network 1, protocol control information such as header information to specify a destination and trailer information to detect errors in the transfer data is added to the

packet. The transmission band instructing part 16 may be designed to make an amendment in consideration of data quantity of such protocol control information. For example, when information of 100 bytes in the sum of the header information and the trailer information is added to data having 1000-byte length, free transfer cycle of only 10% is additionally needed.

[0081] Next, a data transmission operation by the data transmission device 10A will be described below.

[0082] First, the protocol processing part 11 divides transmission data, which is temporarily stored in the data buffer 12, into packets, and adds a header to each divided packet. Such packet generation processing is performed for all the transmission data, and the size of data to be transmitted is instructed to the data transfer part 13. Note that the protocol processing part 11 is practically a CPU (Central Processing Unit) for executing a communications protocol processing program.

[0083] The data buffer 12 temporarily stores the transmission data, and executes operations such as writing of data designated from the protocol processing part 11 and reading of data designated from the data transfer part 13.

[0084] The data transfer part 13 reads only data, which is instructed from the transmission band instructing part 16 and transferred at a time, from the data buffer 12, and subsequently transfers to the data transmission part 14. Such data reading processing is performed for only the data quantity instructed from the protocol processing part 11 in a cycle instructed from the transmission band instructing part 16.

[0085] The data transmission part 14 transmits the transmission data transferred from the data transfer part 13 on the network 1 in synchronization with a clock of the transfer circuit.



(19)

(11) Publication number: 2000253056 A

Generated Document.

PATENT ABSTRACTS OF JAPAN

(21) Application number: 11048595

(51) Intl. Cl.: H04L 12/56 H04L 12/40 H04L 12/28

(22) Application date: 25.02.99

(30) Priority:

(43) Date of application publication: 14.09.00

(84) Designated contracting states:

(71) Applicant: FUJI XEROX CO LTD

(72) Inventor: OZAWA YUJI
HISANAGA TAKANORI

(74) Representative:

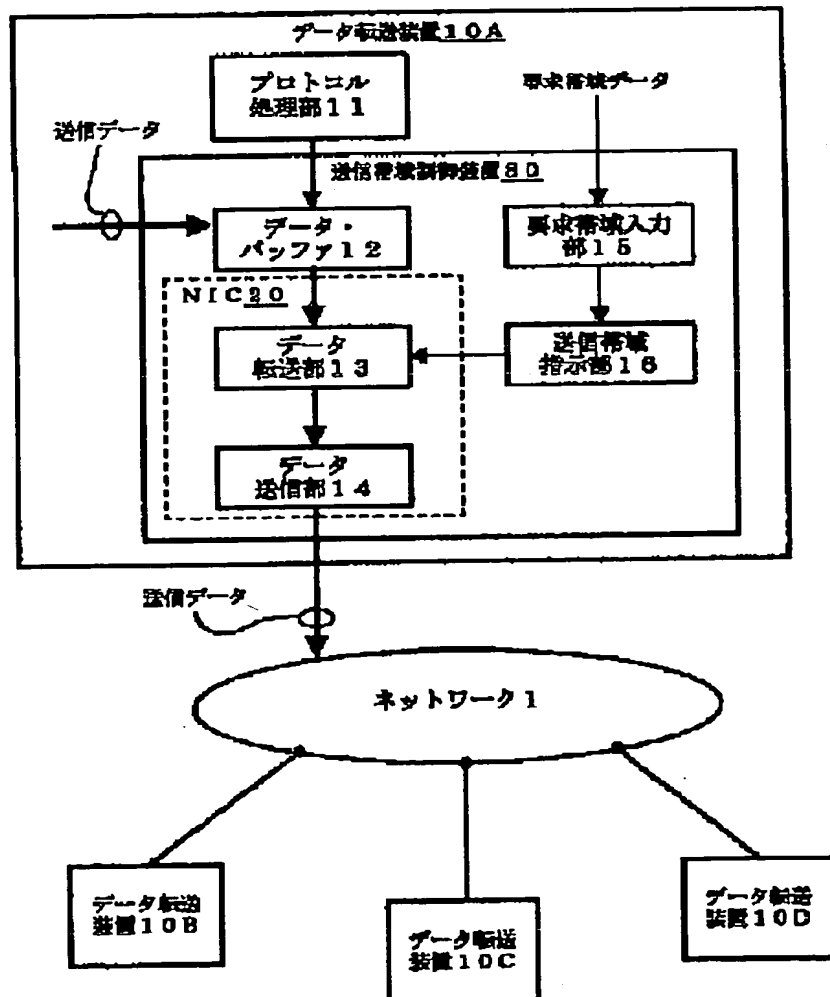
(54) TRANSMISSION BAND CONTROLLER

(57) Abstract:

PROBLEM TO BE SOLVED: To control an oscillation band where a transfer device connected to a network performs data transfer by allowing a data transfer means to calculate a cycle in which data should be transferred from request band information designating data quantity to be transmitted within a certain prescribed time and fixed data quantity and instructing the calculated value to the data transfer means.

SOLUTION: A protocol processing part 11 divides transmission data stored primarily in a data buffer 12 into packets and adds a header to each divided packet. Such packet generation processing is performed for all transmission data and the size of data to be transmitted is instructed to a data transfer part 13. The buffer 12 stores the transmission data temporarily and executes operations such as the writing of data designated from the part 11 and reading of data designated from the part 13. The part 13 reads data quantity which is instructed from a transmission band instructing part 16 and transferred at a time and transfers it.

COPYRIGHT: (C)2000,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-253056

(P2000-253056A)

(43) 公開日 平成12年9月14日 (2000. 9. 14)

(51) Int.Cl.

識別記号

F I

テーマコード(参考)

H 0 4 L 12/56
12/40
12/28H 0 4 L 11/20
11/00
11/201 0 2 B 5 K 0 3 0
3 2 0 5 K 0 3 2
G

審査請求 未請求 請求項の数 8 O L (全 11 頁)

(21) 出願番号

特願平11-48595

(22) 出願日

平成11年2月25日 (1999. 2. 25)

(71) 出願人 000005496

富士ゼロックス株式会社

東京都港区赤坂二丁目17番22号

(72) 発明者 小沢 祐治

神奈川県足柄上郡中井町境430 グリーン

テクノikai 富士ゼロックス株式会社内

(72) 発明者 久永 隆則

神奈川県足柄上郡中井町境430 グリーン

テクノikai 富士ゼロックス株式会社内

(74) 代理人 100086531

弁理士 澤田 俊夫

Fターム(参考) 5K030 HB18 KA03 LC02

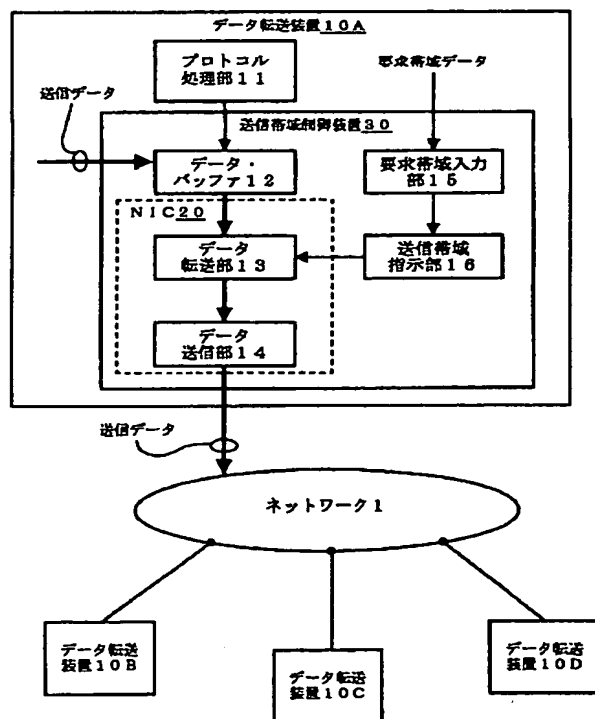
5K032 CC05 DB20

(54) 【発明の名称】 送信帯域制御装置

(57) 【要約】

【課題】 送信データを一時蓄積するバッファを増やすことなく、且つ、ネットワーク・インタフェースの方式に関係なく、要求された送信帯域と実際に送出される送信帯域との誤差をなくすシェーピングを行う。

【解決手段】 要求帯域入力手段は、外部から渡された要求帯域を送信データ量指示手段に通知し、送信データ量指示手段は、この要求帯域に基づいて、一度に転送するデータ量とその転送周期とを計算して、データ転送手段に指示する。データ転送手段は、送信データがデータ・バッファに格納された旨の通知を受け取ると、送信データ量指示手段から指示されたデータ量と転送周期に関する指示に従って、データ・バッファ中の送信データをデータ送信手段に転送する。データ送信手段は、データを受け取るとネットワークに送信する。



【特許請求の範囲】

【請求項 1】 ネットワークに接続されたデータ転送装置における送信帯域を制御するための送信帯域制御装置であって、

送信データを一時記憶するデータ・バッファと、
送信データに所定のプロトコル情報を付加して前記ネットワーク上に送信するデータ送信手段と、

前記データ・バッファと前記データ送信手段間において一定のデータ量毎に所定周期でデータ転送を行なうとともに、データ量と周期を変更可能なデータ転送手段と、ある所定時間内に送信するデータ量を指定した要求帯域情報と、前記一定のデータ量とから、前記データ転送手段がデータ転送すべき周期を算出して、該算出値を前記データ転送手段に指示する送信帯域指示手段と、を具備することを特徴とする送信帯域制御装置。

【請求項 2】 ネットワークに接続されたデータ転送装置における送信帯域を制御するための送信帯域制御装置であって、

送信データを一時記憶するデータ・バッファと、
送信データに所定のプロトコル情報を付加して前記ネットワーク上に送信するデータ送信手段と、

前記データ・バッファと前記データ送信手段間において一定のデータ量毎に所定周期でデータ転送を行なうとともに、データ量と周期を変更可能なデータ転送手段と、ある所定時間内に送信するデータ量を指定した要求帯域情報と、前記所定周期とから、前記データ転送手段がデータ転送すべき一定のデータ量を算出して、該算出値を前記データ転送手段に指示する送信帯域指示手段と、を具備することを特徴とする送信帯域制御装置。

【請求項 3】 ネットワーク接続された他の装置から要求帯域情報を受信するとともに、受信データの有無を前記データ転送手段に通知するデータ受信手段と、
受信データを一時保持するデータ・バッファとをさらに具備し、

前記送信帯域指示手段は、データ・バッファに格納された受信データから要求帯域情報を取り出して、前記データ転送手段に一定のデータ量と周期を指示する、ことを特徴とする請求項 1 又は 2 のいずれかに記載の送信帯域制御装置。

【請求項 4】 前記送信帯域指示手段は、算出した周期に対して、送信するデータ本体とプロトコル制御情報との割合を考慮して補正を行ない、前記データ転送手段に指示することを特徴とする請求項 1 に記載の送信帯域制御装置。

【請求項 5】 前記送信帯域指示手段は、算出した一定のデータ量に対して、送信するデータ本体とプロトコル制御情報との割合を考慮して補正を行ない、前記データ転送手段に指示することを特徴とする請求項 2 に記載の送信帯域制御装置。

【請求項 6】 前記データ転送装置はユーザ入力を受容す

るユーザ・インターフェースを備え、

前記送信帯域指示手段は、前記ユーザ・インターフェースから入力されたデータから要求帯域情報を抽出して、前記データ転送手段に一定のデータ量及び周期を指示することを特徴とする請求項 1 又は 2 のいずれかに記載の送信帯域制御装置。

【請求項 7】 前記データ転送装置はアプリケーションを実行可能であり、

前記送信帯域指示手段は、実行中のアプリケーションから通知されたデータから要求帯域情報を抽出して、前記データ転送手段に一定のデータ量及び周期を指示することを特徴とする請求項 1 又は 2 のいずれかに記載の送信帯域制御装置。

【請求項 8】 前記データ・バッファと前記データ送信手段の間のデータ転送を行なう複数のデータ転送手段と、前記複数のデータ転送手段の各々に対して送信帯域を指示する複数の送信帯域指示手段とを具備することを特徴とする請求項 1 又は 2 のいずれかに記載の送信帯域制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、複数のデータ転送装置がネットワークを共有して行なうデータ転送に係り、特に、ネットワークに接続されたデータ転送装置がデータ転送する送信帯域を制御する送信帯域制御装置に関する。

【0002】 更に詳しくは、本発明は、ネットワーク・インターフェースの方式に関わらず、要求された送信帯域と実際の送信帯域との誤差なくデータ転送を行なうための送信帯域制御装置に係り、特に、送信データを一時蓄積するバッファを増やすことなく、ネットワーク・インターフェースの方式に関係なく、要求された送信帯域と実際に送出される送信帯域との誤差をなくすシェーピングを行うための送信帯域制御装置に関する。

【0003】

【従来の技術】 ネットワーク接続された 2 以上のデータ転送装置 (DTE: Data Terminal Equipment) 間でデータの授受を行う「データ通信」について関する研究開発は、従来から盛んになされてきた。かかるデータ通信の主な意義は、各々のコンピュータ資源の共有や、情報の共有と流通を図ることにある。

【0004】 ここで、ネットワーク接続される各データ転送装置の実体は、専用ハードウェアで構成された通信端末である他、通信ソフトウェアを導入した汎用コンピュータ・システム (例えば、ワークステーションやパーソナル・コンピュータなど) であってもよい。各データ転送装置は、回線終端装置 (DCE: Data Circuit Terminal Equipment) を介して、ネットワークに接続される。

【0005】また、各データ転送装置間を接続するための通信媒体としては、LAN (Local Area Network) のように局所的なものから、2以上のLANどうしを専用線で接続して構成されるWAN (Wide Area Network)、一般公衆電話回線 (PSTN) のように広域的なもの、さらには、各サーバどうしの相互接続の結果として全世界的な巨大ネットワークと化した「インターネット」まで種々様々である。

【0006】通信媒体がPSTNのようなアナログ回線であれば、DCEはモデム (Modulator/Demodulator) である。また、ISDN (Integrated Services Digital Network) であれば、DCEは、TA (ターミナル・アダプタ) とDSU (Digital Service Unit) で構成される。また、LANであれば、DCEはネットワーク・インターフェース・カード (NIC) である。

【0007】NICは、一般には、所謂「アダプタ・カード」の形態で実装され、データ転送を行なうコンピュータ・システムのマザーボード上のバス・スロットに装着して用いられる。

【0008】ネットワークは、通常、多数のデータ通信装置が共用する。ある特定のデータ通信装置間におけるデータ交換のためにネットワークが専有されてしまうと、その間 (例えばデータやメッセージが途切れている間であっても)、他のデータ転送装置は通信を行えなくなってしまう。そこで、複数のデータ転送装置による複数の通信を実現する通信方式として、パケット通信が考案され、広汎に利用されている。

【0009】パケット通信においては、ネットワーク上を通過するデータは、「パケット」と呼ばれる所定長データとして扱われる。パケット交換方式によれば、伝送データは間欠的となるため、ネットワークの共有が可能であり、特定のデータ転送装置間の通信によって伝送回線が長時間専有され続けることはない。パケットは、データ本体の他に、パケットをネットワーク上で正しく移送するためのプロトコル制御情報をヘッダ又はトレイラとして含んでいる。なお、ネットワークを介したデータ通信は、各々のデータ転送装置どうしが所定の通信プロトコルに従った通信処理を行うことにより可能となる。

【0010】ところで、最近では、ネットワークにもQoS (Quality of Service) の考えが導入され始めている。ネットワークにおけるQoSとは、エラーレート、遅延と、帯域の3要素で構成される。このうちの帯域に関して言えば、送信側のデータ転送装置は、ネットワークや受信側が要求する送信帯域に合わせてデータ送信しなければならなくなっている。

【0011】例えば、図6に示すように、同一のサーバ

に対して複数のクライアントが同時にデータ転送を要求するような場合に時間と帯域でスケジューリングを行なうネットワーク・システムでは、通信コストの面から精度良く送信帯域を制御することが必須となってきた。このような送信帯域制御のことを、一般に、「シェーピング」と言う。

【0012】また、図7に示すように、LANからWANに送出するデータ転送のように、物理媒体の帯域が異なるネットワーク間でデータ転送を行なう場合にも、上記のシェーピングが必要である。

【0013】シェーピングの方法としては、クレジット制御方式と、レート制御方式の2方式が当業界で広く知られている。以下、これらの方式について説明する。

【0014】クレジット制御方式とは、「クレジット」、すなわちパケット受信バッファの空き容量に従って送信帯域を制御する方式である。受信側のデータ転送装置は、自身のクレジットを送信側に逐次通知する。送信側のデータ転送装置は、データ・パケットの連続送信数 (すなわち最大ウィンドウ・サイズ) をクレジット以下に制限し、送信パケット分のクレジットを減算することで、パケット受信バッファのサイズを越えるパケットを連続送信しないようにする。

【0015】また、レート制御方式とは、単位時間当たりに送信可能なデータ量を、データ転送開始前あるいはデータ転送中に決めて、送信側では決められたデータ量 (すなわちレート) 以上はデータを送信しないことにより送信帯域を制御する方式である。

【0016】他方、ネットワーク接続されるデータ転送装置は、一般に、汎用コンピュータ・システムを用いて構成される (上述)。この種の汎用コンピュータ・システムは、データ通信処理やその他の各種適用業務処理を実行可能な演算部を備えている。この演算部は、例えば、CPU (Central Processing Unit) という汎用コントローラ・チップで構成される。また、演算部は、通常、実行プログラム・コードをロードしたり作業データを一時格納するためのメモリを、CPUのローカルに配設している。図8には、ネットワーク接続されるデータ転送装置のハードウェア構成を模式的に示している。以下、同図について説明する。

【0017】コンピュータ内部では、演算部とアダプタ・カードその他の周辺装置とは、バスと呼ばれる共通伝送路を介して相互接続されている。

【0018】演算部は、ネットワーク接続するための所定のプロトコル処理を行なうプロトコル処理部と、転送データを一時記憶するためのデータ・バッファを含んでいる。但し、これらは、物理的には、CPUにおいて実行される通信プロトコル・エンティティと、ローカル・メモリ上に割り当てられたプログラム・ロード領域及び転送データ記憶領域を意味するものと理解されたい。

【0019】バスに接続される周辺装置の一例は、ネッ

トワーク接続を果たすためのネットワーク・インターフェース・カード(NIC)である。図示の通り、1つのホスト・コンピュータ上に2以上のNICが装備されていてもよい。

【0020】また、周辺機器の他の例は、ハード・ディスク装置等のディスク型の大容量記憶装置を接続するためのディスク・インターフェース・カードである。当業界では、IDE(Integrated Drive Electronics)やSCSI(Small Computer System Interface) 10 が、標準的なディスク・インターフェース規格である。このディスク・インターフェース・カードのローカルには、ハード・ディスク装置などが配設される。

【0021】図8に示す通り、ネットワーク・インターフェース・カードは、データ転送部と、送信データ・バッファと、データ送信部を含んでいる。データ転送部は、演算部からの送信データの受信、受信データの送信、送信データ・バッファに一時蓄積した送信データの送信など、カード上の処理動作を統括的に制御する。送信データ・バッファは、演算部から送出された送信データを一時蓄積するためのバッファである。また、データ送信部は、送信データ・バッファ中の送信データを、伝送回線のクロックと同期して送出するようになっている。

【0022】このようなハードウェア構成を有するコンピュータ・システムでは、多くの場合、演算部の処理資源を浪費しないように、メモリと周辺装置間のデータ入出力はDMA(Direct Memory Access)方式で行なわれる。例えば、先述したNIC内のデータ転送部は、所謂「バス・マスタ」機能を持ち、メモリ(すなわち演算部内のデータ・バッファ)からNIC内の送信データ・バッファへのデータ転送を、DMA方式、すなわち演算部の介在なしに行なうことができる。DMA転送の間、演算部の処理資源は他の処理に利用される。

【0023】このようなデータ転送装置の機能として、接続されるネットワーク・インターフェースの特性に応じて1度にネットワーク転送可能なデータ・サイズとその転送周期が設定される。

【0024】ここで、上述したホスト・コンピュータのDMA転送機能を参酌した上で、ネットワーク上の送信帯域制御、すなわちシェーピングについて、改めて考察してみる。

【0025】クレジット制御の一例として、TCP(Transmission Control Protocol)プロトコルのフロー制御を利用したシェーピング方法が挙げられる。以下、図9を参照しながら、クレジット制御によるシェーピング方法について説明する。

【0026】送信側のデータ転送装置において、先ず、演算部内のプロトコル処理部では、データ・バッファ部 50

に蓄積されている送信データをバケット単位に分割し、さらに、各バケットにプロトコル制御情報を付加する

(以下では、プロトコル制御情報は「ヘッダ」としてバケットのデータ本体に付加されるものとする)。プロトコル制御情報には、送信データを分割した順番すなわちシーケンス番号が含まれている(周知)。

【0027】次いで、プロトコル処理部は、生成したバケットを、演算部内のデータ・バッファから送信データ・バッファに転送するよう、データ転送部に対して指示する。データ転送部は、DMA転送方式により、指示されたデータ転送を行なう。

【0028】ところで、TCPプロトコルにおけるフロー制御では、確認応答番号によるシーケンス制御を基調としている。すなわち、受信側のデータ転送装置は、バケットを正しく受信できたときにはそのバケットのシーケンス番号(前述)に1だけ加算した確認応答番号を含んだ送達確認(ACK)バケットを返し、一方の送信側のデータ転送装置では、受け取った確認応答番号によってバケットが正しく届いていることを確認する。

【0029】また、逐次的な確認応答番号によるシーケンス制御ではデータ転送が効率的でないため、送信側のデータ転送装置はバケットを幾つかまとめて連続送信し、一方の受信側はまとめて送達確認を行なうという方式も採用されている。送達確認なしに連続送信可能なデータ量は、送信側及び受信側のデータ転送装置間で予め決められており(通常、TCPヘッダ中で指定する)、これを「ウィンドウ・サイズ」と言う。

【0030】送信データ・バッファにバケットが転送されると、データ送信部はバケットを取り出し、これを伝送回線のクロックと同期してネットワークに送出する。このとき、バケットのサイズがネットワーク上において許容されている最大長(MTU(Maximum Transfer Unit)サイズ)を越えている場合には、送信データをさらに分割して送出することもある

(所謂「フラグメンテーション」である。フラグメンテーションは、IP(Internet Protocol)がサポートする)。

【0031】受信側のデータ転送装置では、連続して受信したバケットのエラー・チェックを行う。受信したバケットが正常な場合、受信可能なデータ量の情報を含んだ送達確認(ACK)を、送信側データ転送装置に対して行なう。受信側のデータ転送装置では、このエラー・チェック作業のために、単位時間当たりのデータ転送量を監視して、そのデータ転送量と要求帯域から残データ転送量を算出する機能が必要である。

【0032】他方、送達確認を受信した送信側データ転送装置では、プロトコル処理部は、送達確認(ACK)バケットに含まれている受信可能データ量に応じて、次に送信するバケットの分割及びヘッダ付加を行う。そして、データ転送部に対して次のバケット転送要求を行

う。

【0033】 上述したように、受信側データ転送装置から送信側データ転送装置に通知される送達確認 (ACK) を返す周期と、送信可能なデータ量を制御することにより、シェーピングを行う。

【0034】 次に、レート制御方式によるシェーピング方法について説明する。

【0035】 レート制御方式では、ネットワーク・インタフェースによって予め決められた帯域にシェーピングする。例えば、ATM (Asynchronous Transfer Mode: 非同期転送モード) のように、最小転送データ単位 (ATMの場合は、「セル」と呼ぶ) の送信周期を変えることが可能なネットワークを利用して、データ・フロー (ATMの場合は、VPI (仮想バス識別子) / VCI (仮想チャネル識別子)) 毎に送信バッファを設ける。そして、送信バッファからセルを取り出して、送出する周期を変更することによりシェーピングを実現する。

【0036】 例えば、特開平6-276209号公報には、複数のVP (仮想バス) 又はVC (仮想チャネル) を扱うATM網におけるトラフィック・シェーピング装置について開示されている。

【0037】 このトラフィック・シェーピング装置は、ラインに入力されるセルを送出可能となるまで一時記憶しておくATMセル・メモリと、入力セルをVP又はVCから出力する時刻を判別する出力タイミング判別部と、各時刻毎に出力するセルを記憶する読出制御メモリと、現時刻を示す時計部とを備えている。

【0038】 そして、セル入力時に、ATMセル・メモリに書き込むと同時に、VP又はVCに対応して設定された特性及び履歴により定められる現時刻からそのセルを出力すべき時刻を出力タイミング判別部で判別して、判別された時刻に対応するセルを読み出すためのセル読出情報を読出制御メモリに書き込み、時計部の現時刻情報を用いて読出可能なセルを読出制御メモリで検出し、検出されたセルをATMセル・メモリから読み出すようにしている。

【0039】 このトラフィック・シェーピング装置によれば、ATM網においてセル入力時に予め出力時刻を決定しておくことにより、複数のVP又はVCに対して同時に送出可能かどうかを判別する必要がなく、装置構成を簡素化することができる。

【0040】 この方式について、図10を参照しながら以下に説明する。この方式は、クレジット制御で説明した手順に加え、受信装置から通知されるレート情報が受信されたことを監視し、そのレート情報を送信帯域指示部に通知する要求レート監視部と、通知されたレートを送信周期に換算し、その周期でデータ送受信部を起動する送信帯域指示部をさらに備えることによって、シェーピングを実現している。

【0041】 しかしながら、前者のクレジット制御を利用したシェーピングでは、受信側のデータ転送装置に対して、単位時間当たりのデータ転送量の監視し、そのデータ転送量と要求帯域から残データ転送量を算出する機能を新たに追加するか、あるいは、上記機能を有する別の装置が必要となる。言い換えれば、シェーピング実現のためにコスト高騰は免れない。

【0042】 また、送信側のデータ転送装置は、既存の仕組みを使用するため、通信プロトコル処理をソフトウェアで制御するようにしている。したがって、ハードウェアで制御するような精度良いシェーピング制御は求められない。

【0043】 また、ネットワーク・インタフェースにATMのような送信周期を変更する機能を持たないネットワーク・インタフェース、例えば、Ethernet (米国XEROX社の登録商標) では、シェーピング機能を新たに追加しなければならない。機能追加は、必然的に、コスト増大を招来する。

【0044】 また、後者のレート制御の場合、受信装置から通知されるレート情報が受信されたことを監視し、そのレート情報を送信帯域指示部に通知する要求レート監視部と、通知されたレートを送信周期に換算し、その周期でデータ送受信部を起動する送信帯域指示部をさらに備える必要がある。これ以外に、データ・バス毎に送信帯域値を変えたい場合は、データ・バス毎に送信データを一時的に蓄積するバッファを持つこととなる。すなわち、この場合もコスト高騰は免れない。

【0045】

【発明が解決しようとする課題】 本発明の目的は、ネットワークに接続されたデータ転送装置がデータ転送する送信帯域を制御することができる、優れた送信帯域制御装置を提供することにある。

【0046】 本発明の更なる目的は、ネットワーク・インターフェースの方式に関わらず、要求された送信帯域と実際の送信帯域との誤差なくデータ転送を行なうことができる、優れた送信帯域制御装置を提供することにある。

【0047】 本発明の更なる目的は、送信データを一時蓄積するバッファを増やすことなく、且つ、ネットワーク・インタフェースの方式に関係なく、要求された送信帯域と実際に送出される送信帯域との誤差をなくすシェーピングを行うことができる、優れた送信帯域制御装置を提供することにある。

【0048】

【課題を解決するための手段】 本発明は、上記課題を参酌してなされたものであり、その第1の側面は、ネットワークに接続されたデータ転送装置における送信帯域を制御するための送信帯域制御装置であって、送信データを一時記憶するデータ・バッファと、送信データに所定のプロトコル情報を付加して前記ネットワーク上に送信

するデータ送信手段と、前記データ・バッファと前記データ送信手段間において一定のデータ量毎に所定周期でデータ転送を行なうとともにデータ量と周期を変更可能なデータ転送手段と、ある所定時間内に送信するデータ量を指定した要求帯域情報と前記一定のデータ量とから前記データ転送手段がデータ転送すべき周期を算出して該算出値を前記データ転送手段に指示する送信帯域指示手段とを具備することを特徴とする送信帯域制御装置である。

【0049】また、本発明の第2の側面は、ネットワークに接続されたデータ転送装置における送信帯域を制御するための送信帯域制御装置であって、送信データを一時記憶するデータ・バッファと、送信データに所定のプロトコル情報を付加して前記ネットワーク上に送信するデータ送信手段と、前記データ・バッファと前記データ送信手段間において一定のデータ量毎に所定周期でデータ転送を行なうとともにデータ量と周期を変更可能なデータ転送手段と、ある所定時間内に送信するデータ量を指定した要求帯域情報と前記所定周期とから前記データ転送手段がデータ転送すべき一定のデータ量を算出して該算出値を前記データ転送手段に指示する送信帯域指示手段とを具備することを特徴とする送信帯域制御装置である。

【0050】本発明の第1及び第2の各側面に係る送信帯域制御装置は、ネットワーク接続された他の装置から要求帯域情報を受信するとともに受信データの有無を前記データ転送手段に通知するデータ受信手段と、受信データを一時保持するデータ・バッファとをさらに具備してもよい。また、前記送信帯域指示手段は、データ・バッファに格納された受信データから要求帯域情報を取り出して、前記データ転送手段に一定のデータ量と周期を指示するようにしてもよい。

【0051】また、本発明の第1の側面に係る送信帯域制御装置において、前記送信帯域指示手段は、算出した周期に対して、送信するデータ本体とプロトコル制御情報との割合を考慮して補正を行ない、前記データ転送手段に指示するようにしてもよい。

【0052】また、本発明の第2の側面に係る送信帯域制御装置において、前記送信帯域指示手段は、算出した一定のデータ量に対して、送信するデータ本体とプロトコル制御情報との割合を考慮して補正を行ない、前記データ転送手段に指示するようにしてもよい。

【0053】また、本発明の第1及び第2の各側面に係る送信帯域制御装置において、前記データ転送装置はユーザ入力を受容するユーザ・インターフェースを備え、前記送信帯域指示手段は、前記ユーザ・インターフェースから入力されたデータから要求帯域情報を抽出して、前記データ転送手段に一定のデータ量及び周期を指示するようにしてもよい。データ転送の経路上に帯域の細いネットワークがあることを、ユーザが事前に分かっている場合などに適用することができる。

る場合などに適用することができる。

【0054】また、本発明の第1及び第2の各側面に係る送信帯域制御装置において、前記データ転送装置はアプリケーションを実行可能であり、前記送信帯域指示手段は、実行中のアプリケーションから通知されたデータから要求帯域情報を抽出して、前記データ転送手段に一定のデータ量及び周期を指示するようにしてもよい。データ転送の経路上に帯域の細いネットワークがあることを、アプリケーションが事前に分かっている場合などに適用することができる。

【0055】また、本発明の第1及び第2の各側面に係る送信帯域制御装置において、前記データ・バッファと前記データ送信手段の間のデータ転送を行なうデータ転送手段を複数備えとともに、複数のデータ転送手段の各々に対して送信帯域を指示する送信帯域指示手段を複数備えるようにしてもよい。例えば、データ・バス毎にデータ転送手段及び送信帯域指示手段の組を複数持つことにより、各データ・バスに応じたシェーピングが可能となる。

【0056】

【作用】ネットワーク・インターフェース・カード上に搭載されたデータ転送部は、ネットワークを介したデータ送受信や、カードを受容するホストすなわちデータ転送装置本体側とのデータ送受信など、同カード上における動作を統括的に制御するモジュールである。

【0057】このデータ転送部には、ある一定時間内に転送可能なデータ量と、その周期とを変更するための機能が、従来から備わっている。この変更機能は、本来、汎用性を持たせるために設けられたものである。例えば、ネットワーク・インタフェースが組み込まれるデータ転送装置に、同時に動作する他のインタフェースも組み込まれた場合、両者によるバスの競合を避け、性能を妨げないために設けられたものである。また、上記のデータ転送周期は、バスのクロックを使用するため、正確に刻まれる。

【0058】本発明は、データ転送部のこのような変更機能を利用して、送信データを一時的に蓄積する送信データ・バッファを増やすことなく、要求された送信帯域に誤差なくシェーピングすることを可能とするものである。

【0059】より具体的に言えば、要求帯域入力手段は、外部から渡された要求帯域を、送信データ量指示手段に通知する。送信データ量指示手段は、この要求帯域に基づいて、一度に転送するデータ量とその転送周期とを計算して、データ転送手段に指示する。

【0060】一方、送信データは、プロトコル処理が行なわれた後、データ・バッファに一時格納される。データ転送手段は、送信データがデータ・バッファに格納された旨の通知を受け取る。また、データ転送手段は、送信データ量指示手段から指示されたデータ量と転送周期

に関する指示に従って、データ・バッファ中の送信データをデータ送信手段に転送する。データ送信手段は、データを受け取るとネットワークに送信する。

【0061】以上のような動作手順により、要求された帯域でデータを転送することができる訳である。

【0062】本発明によれば、要求される送信帯域と実際の送信帯域との誤差を、ハードウェア・クロックの最小単位に転送可能なデータ量まで押さえ込むことが可能となる。かかる作用効果は、送信帯域指示手段を追加することにより達成されるものであり、送信データを一時蓄積するためのバッファを増加させる必要がない。

【0063】また、本発明に係る送信帯域指示手段は、ネットワーク・インタフェースとは関係ない箇所への機能追加であるため、どんなネットワーク・インタフェースにも容易に適用可能である。

【0064】本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。

【0065】

【発明の実施の形態】以下、図面を参照しながら本発明の実施例を詳解する。

【0066】図1には、本発明の実施に供されるネットワーク・システム100の構成（第1の実施例）を模式的に示している。ネットワーク・システム100は、データの伝送媒体としてのネットワーク1と、このネットワーク1に接続された2以上のデータ転送装置10A、10B…とで構築される。

【0067】データ転送装置10A…は、ネットワーク・インターフェース・カード（NIC）20を介して、ネットワークに接続される。また、各データ転送装置10A、10B…どうしは、所定の通信プロトコルに従った処理を行うことにより、ネットワーク1を介したデータの送受信が可能となる。かかる通信プロトコルの標準モデルはOSI（Open Systems Interconnection：開放型システム間相互接続）基本参照モデルである。OSI基本参照モデルは、物理層、データリンク層、ネットワーク層、トランスポート層、セッション層、プレゼンテーション層、アプリケーション層という7つの階層で構成される（周知）。ちなみに、前述したTCPプロトコルはトランスポート層に、IPプロトコルはネットワーク層に相当する。

【0068】ネットワーク1に接続される各データ転送装置10A、10B…は、一般に、「ワークステーション」や「パーソナル・コンピュータ」と呼ばれる汎用コンピュータ・システムに通信ソフトウェアを導入して構成される。

【0069】この種の汎用コンピュータ・システムは、データ通信処理やその他の各種適用業務処理を実行可能な演算部の他、実行プログラム・コードをロードしたり作業データを一時格納するためのメモリ、ユーザ・イン

ターフェースとしてのディスプレイ、キーボード及びマウス、外部記憶装置としてのハード・ディスク・ドライブ（HDD）、フロッピー・ディスク・ドライブ（FDD）及びCD-ROMドライブなどを備えている。また、演算部とこれら周辺装置類とは、バスと呼ばれる共通信号伝送路を介して相互接続されている。但し、これらハードウェア構成要素は、当業者には周知であり、且つ、図1は機能実現という観点からデータ転送装置10A…を図解していることから、具体的には図示していない。

【0070】本実施例では、演算部は、所定の通信プロトコル・ソフトウェアを実行するプロトコル処理部11として機能する。

【0071】また、本実施例のデータ転送装置10Aは、ネットワークに接続されたデータ転送装置がデータ転送する送信帯域を制御するための送信帯域制御装置30を含んでいる。この送信帯域制御装置30は、データ・バッファ12と、データ転送部13と、データ送信部14と、要求帯域入力部15と、送信帯域指示部16とを含んでいる。このうち、データ転送部13とデータ送信部14は、図示の通り、ネットワーク・インターフェース・カード（NIC）20上に実装されている。以下、送信帯域制御装置30を構成する各ブロックについて説明する。

【0072】データ・バッファ12は、送信するデータを一時的に蓄積するための装置であり、データ転送部13からの読み出し要求により要求されたデータを出力する。データ・バッファ12の実体は、例えば、演算部（CPU）のローカルに配設されたメモリである。

【0073】データ転送部13は、NIC20内の動作を統括的に制御するための装置である。本実施例のデータ転送部13は、送信帯域指示部16（後述）から通知された転送データ長に応じて、送信データ・バッファ12からデータを読み出し、これを後続のデータ送信部14に転送する。また、このデータ読み出し動作を、送信帯域指示部16から通知されたデータ転送周期に従って、繰り返し回数分だけ実行する。

【0074】データ送信部14は、データ転送部13から転送されたデータにプロトコル制御情報を付加し、伝送回線のクロックに同期してネットワーク1に送信する。

【0075】要求帯域入力部14は、時間内に送出可能なデータ量を特定した要求帯域情報を、送信帯域制御装置30外から受け取って、これを送信帯域指示部16に通知するようになっている。要求帯域情報は、例えば、データ送信部14が送信するための1秒あたりに送信可能なデータ量、あるいは、ネットワーク1上に送出される最小データ単位の送信周期などの形式で記述されている。ここで用いられる要求帯域情報は、システム立ち上げ時から静的に記憶しておいても良い。

7
EX.

【0076】送信帯域指示部16は、要求帯域入力部15から通知された帯域に応じて、1回の転送データ長、データ転送周期と繰り返し回数とを、以下の【数1】に従って算出して、この算出結果をデータ転送部13に通知するようになっている。

【0077】

$$\text{【数1】 } B [\text{Mbps}] = L [\text{bits}] \times 1000000 / R [\mu\text{s}]$$

【0078】ここで、Bは要求帯域、Lはデータ転送部13が1回に転送するデータ量、Rはデータ転送部13がデータを転送する周期を指し、要求帯域はMbps単位で表されるものとする。

【0079】例えば、1回の転送データ量が64ビットの場合に64Mbpsの帯域が要求されると、転送周期は1マイクロ秒となる。

【0080】また、パケットをネットワーク1に送出するときには、送信先を特定するためのヘッダ情報や、送信するデータのエラーを検出するためのトレーラ情報のようなプロトコル制御情報がパケットに付加される。送信帯域指示部16は、このようなプロトコル制御情報のデータ量を考慮した補正を行うようにしてもよい。例えば、1000バイト長のデータに対してヘッダ情報とトレーラ情報が合わせて100バイトだけ付加される場合は、送信周期を10%だけ多く空けるようにすればよい。

【0081】次に、このデータ転送装置10Aによるデータ送信動作について説明する。

【0082】まず、プロトコル処理部11は、データ・バッファ12に一時格納されている送信データをパケットに分割し、分割したパケット毎にヘッダを付加する。このようなパケット生成処理を、全ての送信データ分だけ行ない、送信するデータのサイズをデータ転送部13に指示する。なお、プロトコル処理部11の実体は、通信プロトコル処理プログラムを実行するCPU (Central Processing Unit) であると把握されたい。

【0083】データ・バッファ12は、送信データを一時的に蓄積し、プロトコル処理部11から指定されたデータの書き込み、データ転送部13から指定されたデータの読み出しなどの動作を実行する。

【0084】データ転送部13は、送信帯域指示部16から指示された1回に転送するデータ量だけ、データ・バッファ12から読み出して、後続のデータ送信部14に転送する。また、かかるデータ読み出し処理を、送信帯域指示部16から指示された周期で、プロトコル処理部11から指示されたデータ量分だけ行なう。

【0085】データ送信部14は、データ転送部13から転送されてくる送信データを、伝送回線のクロックに同期して、ネットワーク1上に送出する。

【0086】図2には、送信帯域制御装置30が要求帯

域データを入力するための形態を例示している。既に述べたように、データ転送装置10を構成する汎用コンピュータ・システムは、通常、マウスやキーボード、ディスプレイなどのユーザ・インターフェース17を標準装備している。図2に示す例では、要求帯域は、このユーザ・インターフェース17から通知される要求転送データに基づいて決定する。

【0087】図2に示した方法は、データ転送の経路上に帯域の細いネットワークがあることを、ユーザが事前に分かっている場合などに、適用することができる。

【0088】また、図3には、送信帯域制御装置30が要求帯域データを入力するための他の形態を示している。データ転送装置10を構成する汎用コンピュータ・システムでは、通信プロトコル処理の他、各種アプリケーションが実行されている。この実施形態では、要求帯域データは、プロトコル処理の上位に位置付けられるアプリケーション18から通知される。

【0089】図3に示した方法は、データ転送の経路上に帯域の細いネットワークがあることを、アプリケーション18が事前に分かっている場合などに、適用できる。

【0090】また、図4には、本発明の第2の実施形態に係る送信帯域制御装置30の構成を模式的に示している。

【0091】同図に示すように、この送信帯域制御装置30は、複数のデータ転送部13-1, 13-2, ..., 13-nと、各々のデータ転送部に対応する複数の送信帯域指示部16-1, 16-2, ..., 16-nとを備えている。但し、個々のデータ転送部13-1...及び送信帯域指示部16-1...の構成は、図1に示した第1の実施例と略同一でよい。

【0092】例えば、データ・バス毎にデータ転送部13-1...及び送信帯域指示部16-1...の組を複数持つことにより、各データ・バスに応じたシェーピングが可能となる。この場合、要求帯域指示部15は、入力した要求帯域データがどのデータ・バスへの帯域指示かを判別して、そのデータ・バスに対応する送信帯域指示部16...に分配すればよい。そして、送信帯域指示部16...は、図1で示した第1の実施例における場合と同様に、データ量と周期を算出して、該当するデータ転送部13...に算出結果を通知すればよい。

【0093】データ・バッファ12やデータ送信部14は、上述した第1の実施例と同様の構成及び動作特性を有することにより、複数のデータバスに対応することができる。

【0094】また、図5には、本発明の第3の実施形態に係る送信帯域制御装置30の構成を模式的に示している。

【0095】図5に示す通り、この実施例では、要求帯域データは、ネットワーク1を介して接続される他のデ

ータ転送装置から通知される。ここで言うデータ転送装置は、具体的には受信側のデータ転送装置のことである。

【0096】ネットワーク、または、ネットワークに接続された他のデータ転送装置から通知される要求帯域データは、データ受信部19によって受信される。データ受信部19は、要求帯域データを受信したことを、データ転送部13に通知する。

【0097】データ転送部13は、通知された要求帯域データを、データ・バッファ12に転送する。

【0098】プロトコル処理部11は、データ・バッファ12に転送された要求帯域データを解析し、要求帯域を抽出し、要求帯域入力部15に通知する。

【0099】要求帯域入力部15及び送信帯域指示部16は、上述した第1の実施例と同様に動作する。すなわち、送信帯域指示部16は、データ量と周期を算出して、データ転送部13に算出結果する。

【0100】[追補] 以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参酌すべきである。

【0101】

【発明の効果】以上詳記したように、本発明によれば、ネットワークに接続されたデータ転送装置がデータ転送する送信帯域を制御することができる、優れた送信帯域制御装置を提供することができる。

【0102】また、本発明によれば、ネットワーク・インターフェースの方式に関わらず、要求された送信帯域と実際の送信帯域との誤差なくデータ転送を行なうことができる、優れた送信帯域制御装置を提供することができる。

【0103】また、本発明によれば、送信データを一時蓄積するバッファを増やすことなく、且つ、ネットワーク・インターフェースの方式に関係なく、要求された送信帯域と実際に送出される送信帯域との誤差をなくすシェーピングを行うことができる、優れた送信帯域制御装置を提供することができる。

【0104】本発明によれば、要求される送信帯域と実際の送信帯域との誤差を、ハードウェア・クロックの最

小単位に転送可能なデータ量まで押さえ込むことが可能となる。かかる作用効果は、送信帯域指示部を追加することにより達成されるものであり、送信データを一時蓄積するためのバッファを増加させる必要がない。また、本発明に係る送信帯域指示部は、ネットワーク・インターフェースとは関係ない箇所への機能追加であるため、どんなネットワーク・インターフェースにも容易に適用可能である。

【図面の簡単な説明】

10 【図1】 本発明の実施に供されるネットワーク・システム100の構成を模式的に示した図である。

【図2】 送信帯域制御のための要求帯域情報をユーザ・インターフェースを介して受け取る構成を概念的に示した図である。

【図3】 送信帯域制御のための要求帯域をアプリケーションから受け取る構成を概念的に示した図である。

【図4】 本発明の第2の実施例に係る送信帯域制御装置の構成を模式的に示した図である。

20 【図5】 本発明の第3の実施例に係る送信帯域制御装置の構成を模式的に示した図である。

【図6】 ネットワーク・システムの概略構成（従来例）を示した図である。

【図7】 複数のネットワークを介してデータ転送される様子を模式的に示した図である。

【図8】 従来のデータ転送装置の概略構成を示した図である。

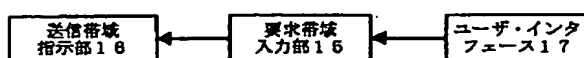
30 【図9】 TCP (Transmission Control Protocol) を利用したクレジット制御の動作シーケンス（従来例）を模式的に示した図である。

【図10】 ATM (Asynchronous Transfer Mode: 非同期通信モード) を利用してレート制御を行なうデータ転送装置の概略構成を示した図である。

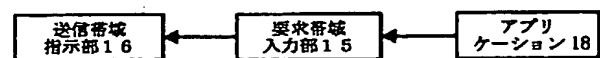
【符号の説明】

1…ネットワーク、10…データ転送装置、12…データ・バッファ、13…データ転送部、14…データ送信部、15…要求帯域入力部、16…送信帯域指示部、17…ユーザ・インターフェース、18…アプリケーション、データ受信部19、20…ネットワーク・インターフェース・カード、30…送信帯域制御装置、100…ネットワーク・システム。

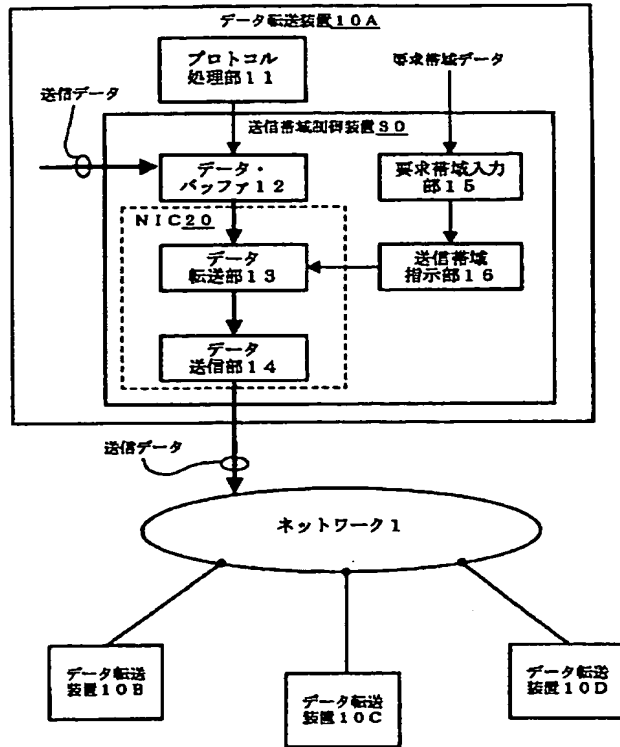
【図2】



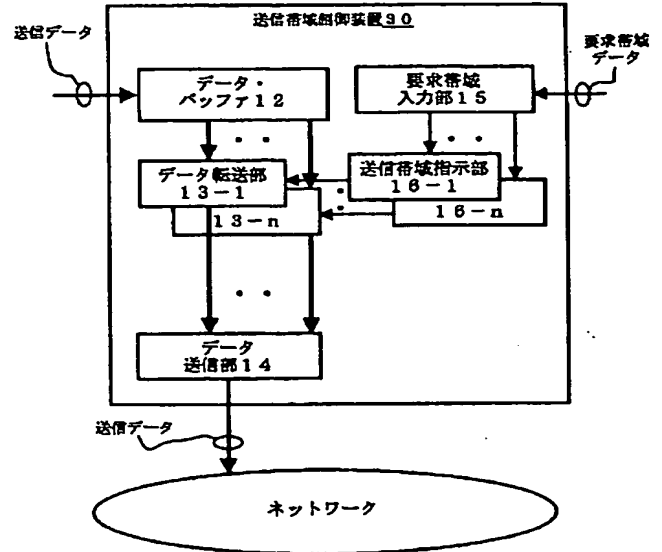
【図3】



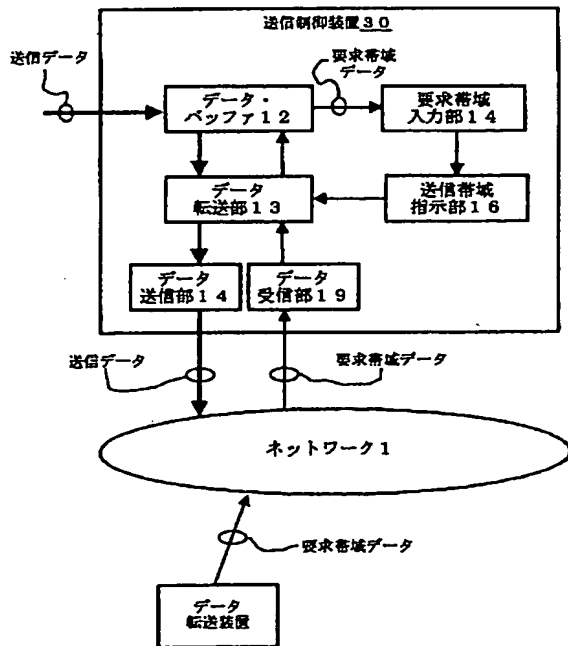
【図1】



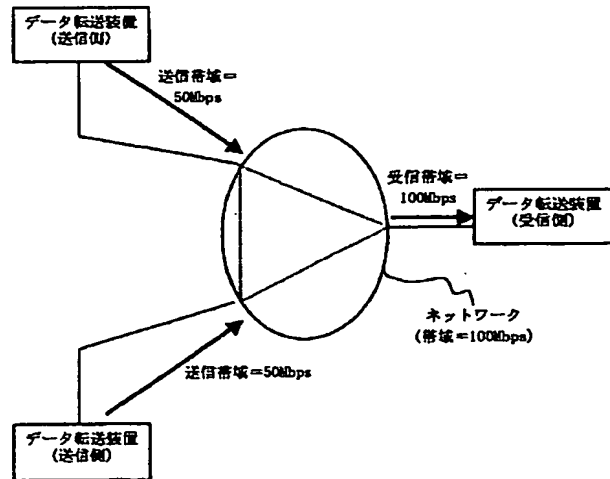
【図4】



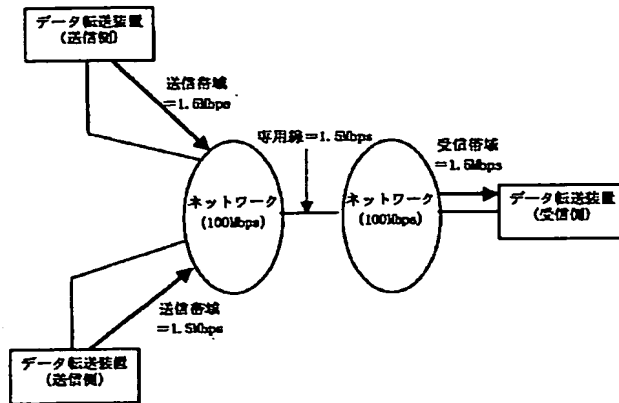
【図5】



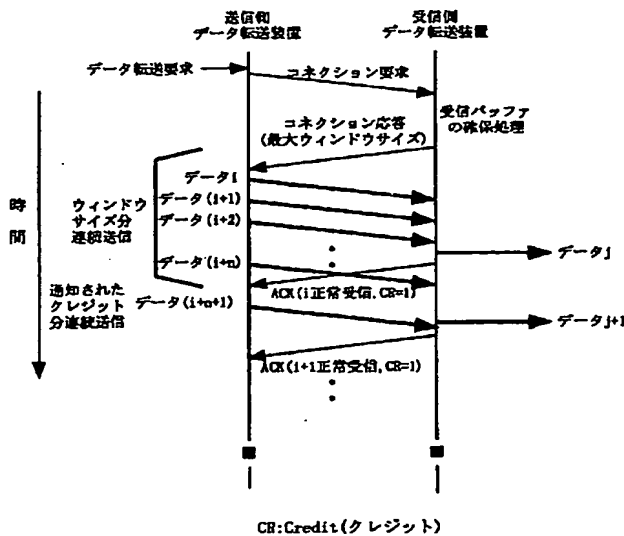
【図6】



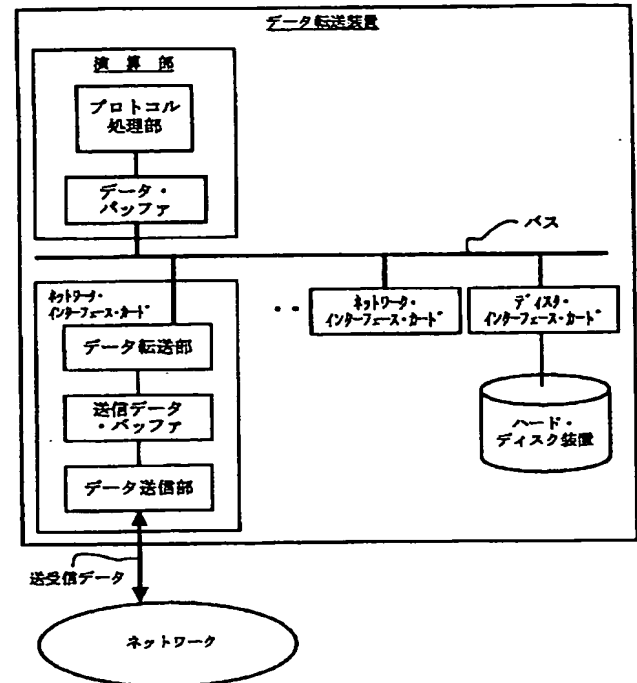
【図7】



【図9】



【図8】



【図10】

